Between Transfinite and Dynamical Proof Methods

Peter Schuster

joint work with Daniel Misselbeck–Wessel and partly with Giulio Fellin

Dipartimento di Informatica, Università di Verona

Workshop in Honour of Thierry Coquand's 60th Birthday

Göteborg, 24–26 August 2022

Abstract

The dynamical proof method was developed to make sense in constructive mathematics of the numerous instances of the axiom of choice modern algebra abounds with. Recently the dynamical method has proved to reach into general algebra, and thus to cover more universal transfinite proof principles. This prompts the question whether dynamical methods vindicate their transfinite forerunners a century after, on top of substituting them well for computation.

Technically, a key role is played by an inductive generalisation of the Jacobson radical membership to which amounts to termination of generation trees. As for dynamical algebra proper, the ideal objects postulated by the transfinite methods are approximated by paths in finite binary trees which proceed at branchings as if they actually were ideal objects.

Computational content

"When we say that we have a constructive version of an abstract algebraic theorem, this means that we have a theorem
▷ the proof of which is constructive,
▷ which has a clear computational content, and
▷ from which we can recover the usual version of the abstract theorem by an immediate application of a well classified non-constructive principle . . . "

Thierry Coquand, Henri Lombardi, "Hidden constructions in abstract algebra: Krull dimension of distributive lattices and commutative rings", 2002

Disclaimer

For staying closer to common practice we work in **IZF**, with heuristics in **ZF(C)**.
With some effort a predicative treatment, e.g. in **CZF**, should be possible.

Prototype

The *Jacobson radical* $\mathrm{Jac}(J)$ of an ideal J of a commutative ring A with 1 is usually defined as the intersection of the maximal ideals of A which contain J:

$$\mathrm{Jac}(J) = \bigcap Max(A)/J = \{a \in A \mid \forall M \in Max(A)\,(J \subseteq M \Rightarrow a \in M)\} \quad (1)$$

where an ideal M of A is *maximal* if it is such among the *proper* ideals ($\neq A$).

With the *Axiom of Choice* (AC), the definition of $\mathrm{Jac}(J)$ is equivalent to

$$\mathrm{Jac}(J) = \{a \in A \mid \forall b \in A\,(\langle a, b\rangle \ni 1 \Rightarrow \langle J, b\rangle \ni 1)\} \quad (2)$$

where angle brackets $\langle \ldots \rangle$ denote generated ideals.

In constructive algebra, the *first-order* alternative (2) is taken as the definition of $\mathrm{Jac}(J)$, so as to provide for a computationally meaningful concept.

With AC at hand the *second-order* original (1) then becomes a theorem, which we henceforth refer to as the *Jacobson Intersection Principle* (JIP).

With classical logic, in fact, JIP is equivalent to Krull's *Maximal Ideal Theorem* (MIT), tantamount to AC: *every proper ideal is contained in a maximal ideal*.

Prime, maximal and ultra

By contrast, the *Boolean Ultrafilter Theorem* (BUT) suffices for Krull's *Prime Ideal Theorem* (PIT): *every proper ideal is contained in a prime ideal*.

With classical logic every maximal ideal is prime, but not always conversely.

Correspondingly, MIT is stronger than PIT—just as AC is stronger than BUT.

Prime and maximal ideals are the same thing in Boolean rings where $a^2 = a$, and thus in Boolean algebras. In the latter the dual concepts of prime and maximal filters both coincide with the concept of ultrafilters.

In a commutative ring, a proper ideal I is *prime* if

$$a \times b \in I \Longrightarrow a \in I \vee b \in I$$

For distributive lattices, replace multiplication $\times$ by meet $\wedge$.

For filters rather than ideals, dualize by swapping $\wedge$ and $\vee$.

In a Boolean algebra, a proper filter F is an *ultrafilter* if

$$a \in F \vee \neg a \in F$$

The ultrafilters of the Lindenbaum algebra are the *complete consistent theories*.

Challenge

As we have dealt with PIT elsewhere, we now focus on maximal ideals etc.

Logically speaking, MIT is *model existence* and JIP is *semantic conservation*. Shifting focus from the former to the latter has helped us to pin down the computational content of MIT as the *syntactical conservation* underlying JIP.

Can we generalise, abstracting from the (relatively rich) setting of rings?

Do also maximality principles conceptually closer to AC, such as the Teichmüller–Tukey Lemma (TTL), have a syntactical underpinning?

We will first solve semantically and then interpret syntactically the problem

Which principle X is to TTL just as JIP is to MIT? (3)

Heuristics comes from further cases: the Jacobson radical can also be defined

— for ideals J of distributive lattices D [Coquand '03, Coquand–Lombardi '02]

$$\mathrm{Jac}(J) = \{\, a \in D \mid \forall b \in D\,(a \vee b = 1 \Rightarrow J \vee b = 1)\,\}$$

— for theories Γ of propositional languages P

$$\mathrm{Jac}(\Gamma) = \{\, \alpha \in P \mid \forall \beta \in P\,(\alpha, \beta \vdash \bot \Rightarrow \Gamma, \beta \vdash \bot)\,\}$$

The Jacobson radical of a propositional theory (with G. Fellin)

Let Γ be a theory of intuitionistic provability $\vdash_i$ in a propositional language P, where by a theory we understand a set of formulas which is deductively closed.

The Jacobson radical of Γ boils down to the *stable closure* of Γ:

$$\begin{aligned}\mathrm{Jac}(\Gamma) &= \{\, \alpha \in P \mid \forall \beta \in P\, (\alpha, \beta \vdash_i \bot \Rightarrow \Gamma, \beta \vdash_i \bot)\,\} \\ &= \{\, \alpha \in P \mid \forall \beta \in P\, (\alpha \vdash_i \neg\beta \Rightarrow \Gamma \vdash_i \neg\beta)\,\} \\ &= \{\, \alpha \in P \mid \Gamma \vdash_i \neg\neg\alpha)\,\}\end{aligned}$$

The (maximal) ideals correspond in logic to the (complete consistent) theories.

Hence JIP instantiates to a slight variant of *Lindenbaum's Lemma*:

$$\mathrm{Jac}(\Gamma) = \bigcap \{\, \Theta \mid \Theta \text{ complete consistent theory, } \Theta \supseteq \Gamma\,\}$$

The syntactical underpinning of this is well known as *Glivenko's Theorem*:

$$\Gamma \vdash_i \neg\neg\alpha \quad \Longleftarrow \quad \Gamma \vdash_c \alpha$$

As all consistent complete theories are stable under $\neg\neg$, for them it is irrelevant whether deductive closure is understood for $\vdash_i$ or for classical provability $\vdash_c$.

Setting

We abstract from ideals and theories to the elements of a complete lattice L, and from comaximality $\langle \ldots \rangle \ni 1$ and inconsistency $\ldots \vdash \bot$ to a *Scott-open* subset O of L: that is, a monotone predicate that splits directed joins.

We assume that $1 \in O$, and sometimes that O be a Scott-open *filter* of L, i.e. meet-closed as well. — Think of $O(x)$ as "x generates 1" or "x proves $\bot$".

The *generalised radical* $j : L \to L$ defined by $jx = \bigvee J_x$ where

$$J_x = \{\, a \in L \mid \forall b \in L\,(O(a \vee b) \Rightarrow O(x \vee b))\,\}$$

is a closure operator generalising the aforementioned Jacobson radicals.

Some key features of j are that j is the largest closure operator on L for which $O = j^{-1}(1)$; and that if L is distributive, then O is a filter iff j is a nucleus.

In the vein of logical theories, we say that $x \in L$ is *complete* if for every $y \in L$

$$\text{either } y \leqslant x \text{ or } O(y \vee x)$$

This is equivalent to x being *maximal* in the sense that for every $y \in L$

$$\text{if } x \leqslant y, \text{ then either } x = y \text{ or } O(y)$$

We further say that $z \in L$ is *proper* or *consistent* if $\neg O(z)$.

Achievements

With classical logic and AC we can prove the following for every $x \in L$:

(a) *the radical jx is the meet of all proper complete $y \geqslant x$;*

(b) *if x is proper, then there is a proper complete $y \geqslant x$.*

As the proper complete ideals are just the maximal ideals, (a) and (b) generalise JIP and MIT, respectively. If L is algebraic, then (b) generalises TTL.

With (a) we thus obtain the desired semantic solution X of (3), and can search its syntactical interpretation. To this end we put (b) "upside up":

(c) *any given $x \in L$ belongs to O if(f) every complete $y \geqslant x$ is in O.*

Adapting our recent syntactical treatments of PIT and of some specific maximality principles, we inductively define a collection T of finite binary trees labelled by elements of L, with an appropriate termination concept for paths.

All this allows us to prove constructively—in particular, without AC—the following syntactical counterpart of (c) whenever O is a filter:

(d) *any given $x \in L$ belongs to O if(f) there is a tree in T with root labelled by x such that every branch of the tree terminates in O.*

This is reminiscent of weeping willows: the tip of every twig dips into O.

weeping willow/saule pleureure/tårpil/Trauerweide/salice piangente

Inductive generation

For each $a \in L$ we name $\overline{a}$ the pseudo-complement of a with respect to O:

$$\overline{a} = \{b \in L : O(a \vee b)\}$$

Any $a \in L$ is complete precisely when, for every $b \in L$, either $b \leqslant a$ or $b \in \overline{a}$.

Definition

We generate the binary relation $\sqsubseteq$ on L inductively by the following three rules:

$$\frac{x \leqslant y}{x \sqsubseteq y} \qquad \frac{O(y)}{x \sqsubseteq y} \qquad \frac{x \sqsubseteq y \vee a \quad \forall b \in \overline{a}\,(x \sqsubseteq y \vee b)}{x \sqsubseteq y}$$

Proposition (**IZF**)

$x \sqsubseteq y$ is equivalent to $x \leqslant jy$.

Corollary (**IZF**)

$jy = \bigvee \{\, x \in L \mid x \sqsubseteq y \,\}$

We could stop here . . . but some of us grew up in the woods . . .

Trees

Working with the pseudocomplement $\overline{a}$ allows us to get by with binary trees rather than infinitely branching ones.

Definition

We inductively define a collection T_x of finite labelled binary trees t over $x \in L$:

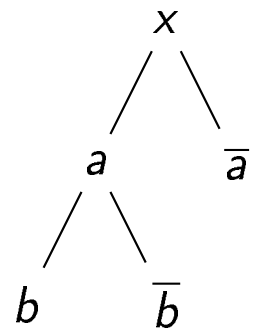
Base The root-only tree labelled with x belongs to T_x.

Step If $t \in T_x$ has a path π, and $a \in L$, then let the leaf of π have two children labelled by a and $\overline{a}$. Add the extended tree to T_x.

Mind that paths lead from the root of a tree to one of its leaves.

Set $T = \bigcup\{T_x : x \in L\}$.

E.g., if $x, a, b \in L$, then the following tree belongs to T_x:



Termination

Next, we define a relation $\vdash$ between paths π, represented as strings of elements of $L \cup \{\, \overline{a} \mid a \in L \,\}$, and elements m of L recursively over π as follows:

$$\begin{aligned} x \vdash m &\equiv m \leqslant jx \\ x\pi a \vdash m &\equiv (x \vee a)\pi \vdash m \\ x\pi\overline{a} \vdash m &\equiv \forall b \in \overline{a}\,(x\pi b \vdash m). \end{aligned}$$

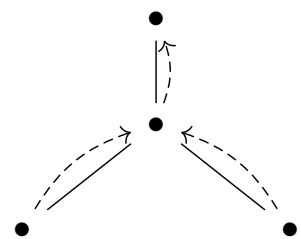
Read $\pi \vdash m$ as "the data accumulated along path π yields information m"

By unfolding the inductive definition of T_x the intended becomes evident:

Lemma (IZF)

Let π be a path of some tree $t \in T_x$, and let $a, m, m' \in L$. Then:

$$\frac{\pi a \vdash m \qquad \pi\overline{a} \vdash m'}{\pi \vdash m \wedge m'}$$



Definition

Let $M \subseteq L$. We say that a tree $t \in T_x$ *terminates in* M if for every path π of t there is $m \in M$ such that $\pi \vdash m$.

Read termination as "every possible run of the program yields a desired result"

Folding up

By the Lemma one can easily climb any such tree from the leaves to the root:

Theorem (**IZF**)

Let $M \subseteq L$ be meet-closed. For every $x \in L$, the following are equivalent.

1. *There is $m \in M$ such that $m \leqslant jx$, i.e. $x \vdash m$.*
2. *The root-only tree labelled with x terminates in M.*
3. *There is a tree $t \in T_x$ which terminates in M.*

In other words, membership in a generalised radical amounts to termination.

The case of Scott-open filters $M = O$ gives us the desired result (d):

Corollary (**IZF**)

Suppose that O is a Scott-open filter. For every $x \in L$ we have $O(x)$ if and only if there is a tree $t \in T_x$ which terminates in O.

Dynamical methods

Remember the pseudo-complement $\overline{a}$ of $a \in L$ with respect to O:

$$\overline{a} = \{b \in L : O(a \vee b)\}$$

Also, $a \in L$ is complete precisely when, for every $b \in L$, either $b \leqslant a$ or $b \in \overline{a}$.

Every tree $t \in T_x$ represents the course of a dynamic argument *as if* we would have a complete $c \geqslant x$.

Every complete $c \geqslant x$ gives indeed rise to a path through t: at each branching, corresponding to some $a \in L$, by completeness either $a \leqslant c$ or $c \in \overline{a}$.

Towards the Teichmüller–Tukey Lemma (TTL)

We say that a subset O of L is *of cofinite character* if, for every $x \in L$,

$$x \in O \quad \text{iff} \quad Kx \between O$$

where $Kx = \{\, a \in KL \mid a \leqslant x \,\}$ and KL is the set of compact elements of L.

In algebraic lattices the Scott-open subsets are precisely those of cofinite character, and the radical anyway is an affair of compact elements only:

$$x \leqslant jy \quad \text{iff} \quad \forall a \in KL\,(O(x \vee a) \Rightarrow O(y \vee a))$$

Example (**ZFC**)

Let S be a set. Then $L = \mathrm{Pow}(S)$ is an algebraic lattice with $KL = \mathrm{Fin}(S)$.

Recall that $F \subseteq \mathrm{Pow}(S)$ is of *finite character* whenever

$$T \in F \quad \text{iff} \quad \mathrm{Fin}(T) \subseteq F$$

for every $T \subseteq S$. Under this condition upon F, the original form of TTL reads:

every element of F can be extended to a maximal one.

Instead one can work directly with the complement of cofinite character:

$$O = \mathrm{Pow}(S) \setminus F$$

Abstract dependence

Let $\rhd$ be a consequence relation with the *Mac Lane-Steinitz exchange property*:

$$\text{if } U, a \rhd b, \text{ then either } U \rhd b \text{ or } U, b \rhd a\,.$$

A subset T of S is *dependent*, for short $\mathrm{dep}(T)$, if there is $a \in T$ along with $U \in \mathrm{Fin}(T - \{a\})$ such that $U \rhd a$. Independence is to be non-dependence.

A *basis* for S is an independent subset T which is *spanning*: that is, $\langle T \rangle = S$. For example, $\emptyset$ is independent and S is spanning; and $\mathrm{dep}(\{a\})$ iff $\emptyset \rhd a$.

Dependence is of cofinite character:

$$O(T) \equiv \mathrm{dep}(T).$$

The proper complete $T \subseteq S$ are, in **ZF**, precisely the bases for S. So, in **ZFC**:

$$jT = \bigcap \mathrm{Bases}/T.$$

Alternatively, in rather finite terms:

$$a \in jT \quad \text{iff} \quad \forall U \in \mathrm{Fin}(S)\,(\mathrm{dep}(a, U) \Rightarrow \mathrm{dep}(T, U))\,.$$

In particular, $jT = S$ for every dependent $T \subseteq S$.

Linear dependence

Let K be a discrete field of characteristic $\neq 2$, and S a discrete K-vector space. Let $\rhd$ on S be given by linear span:

$$a_1, \ldots, a_k \rhd b \equiv \exists \lambda_1, \ldots, \lambda_k \in \mathsf{K} \sum_{i=1}^{k} \lambda_i a_i = b,$$

In this prime example of a consequence relation with the exchange property, the terms (in)dependence, spanning and basis have their customary meanings.

Every independent $T \subseteq S$ is *radical*: that is, $jT = T$. Hence, in **ZFC**:

$$T = \bigcap \mathrm{Bases}/T$$

Simply because $0 \notin T$, this yields Hamel's theorem from 1910:

every independent subset T can be extended to a basis for S.

With TTL, by contrast, the independent subsets above T form $\mathcal{F} \subseteq \mathsf{Pow}(S)$ of finite character; and every maximal independent subset of S is spanning.